

**PATENT APPLICATION**

**TECHNIQUES FOR USING A CAPTURED IMAGE FOR THE  
RETRIEVAL OF RECORDED INFORMATION**

Inventor(s): Berna Erol, a citizen of Canada, residing at  
1422 Bellevue Avenue, Unit 105  
Burlingame, CA 94010

Jonathan J. Hull, a citizen of The United States, residing at  
751 Laurel Street, PMB 434  
San Carlos, CA 94070

Assignee: Ricoh Company, Ltd.  
3-6 Naka-magome, 1-Chome, Ohta-ku  
Tokyo, Japan 143

Entity: Large

## **TECHNIQUES FOR USING A CAPTURED IMAGE FOR THE RETRIEVAL OF RECORDED INFORMATION**

### **CROSS-REFERENCES TO RELATED APPLICATIONS**

- 5    [0001]    The present application is a non-provisional of and claims priority from U.S. Provisional Application No. 60/462,412, Attorney Docket No. 15358-009400US, filed April 11, 2003, the entire contents of which are herein incorporated by reference for all purposes.
- [0002]    The present application incorporates by reference for all purposes the entire contents of the following:
- 10   [0003]    U.S. Application No. 10/412,757, Attorney Docket No. Docket No. 15358-009500US, filed April 11, 2003;
- [0004]    U.S. Application No. \_\_/\_\_,\_\_, Attorney Docket No. 15358-009410US, filed September 12, 2003;
- [0005]    U.S. Application No. \_\_/\_\_,\_\_, Attorney Docket No. 15358-009420US, filed  
15   September 12, 2003;
- [0006]    U.S. Application No. 10/660,867, Attorney Docket No. 15358-009430US, filed September 12, 2003; and
- [0007]    U.S. Application No. 10/001,895, Attorney Docket No. 15358-006500US, filed November 19, 2001.

20

### **BACKGROUND OF THE INVENTION**

- [0008]    The present application generally relates to techniques for accessing recorded information and more particularly to techniques for accessing recorded information using an input image.
- 25   [0009]    Recording information during presentations has gained a lot of popularity in recent years. For example, colleges and universities have started to program classes and lectures, corporations have started to record meetings and conferences, etc. The information during a presentation may be recorded using one or more capture devices. The recorded information

may comprise different types or streams of information including audio information, video information, and the like.

[0010] The recorded information is then available for use by a user after the presentation. The conventional way for accessing these recordings has been by viewing the recordings sequentially. More efficient techniques are desired for accessing or retrieving the recorded information or indexing into the recorded information.

## BRIEF SUMMARY OF THE INVENTION

[0011] Embodiments of the present invention relate to determining portions of recorded information using an input image.

[0012] In one embodiment, a method for determining a recorded presentation information document is provided. The method comprises: receiving information identifying an input image; comparing the input image with a plurality of image file documents to determine an image file document in the plurality of image file documents that includes information that is considered to match the input image; and determining a recorded presentation information document that is associated with the image file document that was determined.

[0013] In another embodiment, a method for determining a recorded presentation information document is provided. The method comprises: determining a captured image, the captured image including a display; determining contents of the display; and using the contents of the display to search a plurality of recorded images documents to identify one or more recorded images documents that include the contents.

[0014] A further understanding of the nature and advantages of the invention herein may be realized by reference of the remaining portions in the specifications and the attached drawings.

## BRIEF DESCRIPTION OF THE DRAWINGS

[0015] Fig. 1 is a simplified block diagram of a system that may incorporate an embodiment of the present invention;

[0016] Fig. 2 depicts a simplified flow chart for determining and storing association information for image file according to one embodiment of the present invention;

[0017] Fig. 3 depicts a simplified flow chart for a method of determining a portion of a recorded presentation information document according to one embodiment of the present invention;

[0018] Fig. 4 depicts a simplified flow chart of an alternate embodiment for determining a portion in a recorded presentation information document using an input image according to embodiments of the present invention;

[0019] Fig. 5 depicts an interface that may be used to specify an input image according to one embodiment of the present invention;

[0020] Fig. 6 depicts an interface that displays a portions of the recorded presentation information document in response to receiving an input image according to one embodiment of the present invention;

[0021] Fig. 7 depicts a simplified flow chart of a method for using an input image to perform a search according to one embodiment of the present invention;

[0022] Fig. 8 depicts a simplified flowchart of a method for determining television programs using an input image according to one embodiment of the present invention;

[0023] Fig. 9 depicts an embodiment showing using a digital camera for retrieving a collection of recorded presentation information documents and determining a portion in the recorded presentation information documents according to embodiments of the present invention;

[0024] Fig. 10 depicts a simplified flowchart of a method for performing a disambiguation process in determining a portion of one or more recorded presentation information documents according to one embodiment of the present invention; and

[0025] Fig. 11 is a simplified block diagram of a data processing system that may incorporate an embodiment of the present invention;

## DETAILED DESCRIPTION OF THE INVENTION

[0026] In the following description, for the purposes of explanation, specific details are set forth in order to provide a thorough understanding of the invention. However, it will be apparent that the invention may be practiced without these specific details.

[0027] Fig. 1 is a simplified block diagram of a system 100 that may incorporate an embodiment of the present invention. System 100 depicted in Fig. 1 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize other variations, modifications, and alternatives.

[0028] System 100 includes a computer system 102 that may be used by a user to display information at a presentation. Examples of presentations include lectures, meetings, conferences, classes, speeches, demonstrations, etc. The presentation material may include slides, photos, audio messages, video clips, text information, web pages, etc. The user may use one or more applications 104 executed by computer 102 to generate the presentation material. An example of a commonly used application for preparing slides to be presented at a presentation is PowerPoint™ provided by Microsoft™ Corporation. For example, as depicted in Fig. 1, the user may use PowerPoint™ application 104 to create a "presentation.ppt" file 106 (\*.ppt file). A \*.ppt file created using a PowerPoint™ application may comprise one or more pages, each page comprising one or more slides. A \*.ppt file may also store information as to the order in which the slides are to be presented at the presentation and the manner in which the slides will be presented.

[0029] In addition to PowerPoint™ presentation files comprising slides, other types of files comprising other presentation material may also be created using different applications executed by computer 102. These files may be referred to in general as "symbolic presentation files". A symbolic presentation file is any file created using an application or program and that comprises at least some content that is to be presented or output during a presentation. A symbolic presentation file may comprise various types of contents such as slides, photos, audio messages, video clips, text, web pages, images, etc. A \*.ppt file created using a PowerPoint™ application is an example of a symbolic presentation file that comprises slides.

[0030] Capture devices 118 are configured to capture information presented at a presentation. Various different types of information output during a presentation may be captured or recorded by capture devices 118 including audio information, video information, images of slides or photos, whiteboard information, text information, and the like. For purposes of this application, the term "presented" is intended to include displayed, output, spoken, etc. For purposes of this application, the term "capture device" is intended to refer to

any device, system, apparatus, or application that is configured to capture or record information of one or more types. Examples of capture devices 118 include microphones, video cameras, cameras (both digital and analog), scanners, presentation recorders, screen capture devices (e.g., a whiteboard information capture device), symbolic information capture  
5 devices, etc. In addition to capturing the information, capture devices 118 may also be able to capture temporal information associated with the captured information.

**[0031]** A presentation recorder is a device that is able to capture information presented during a presentation, for example, by tapping into and capturing streams of information from an information source. For example, if a computer executing a PowerPoint™ application is  
10 used to display slides from a \*.ppt file, a presentation recorder may be configured to tap into the video output of the computer and capture keyframes every time a significant difference is detected between displayed video keyframes of the slides. The presentation recorder is also able to capture other types of information such as audio information, video information, slides information stream, etc. Temporal information may also be captured. The temporal  
15 information associated with the captured information indicating when the information was output or captured is then used to synchronize the different types of captured information. Examples of presentation recorders include a screen capture software application, a PowerPoint™ application that allows recording of slides and time elapsed for each slide during a presentation, presentation records described in U.S. Application No. 09/728,560,  
20 filed November 30, 2000 (Attorney Docket No. 15358-006210US), U.S. Application No. 09/728,453, filed November 30, 2000 (Attorney Docket No. 15358-006220US), and U.S. Application No. 09/521,252, filed March 8, 2000 (Attorney Docket No. 15358-006300US).

**[0032]** A symbolic information capture device is able to capture information stored in symbolic presentation documents that may be output during a presentation. For example, a  
25 symbolic information capture device is able to record slides presented at a presentation as a sequence of images (e.g., as JPEGs, BMPs, etc.). A symbolic information capture device may also be configured to extract the text content of the slides. For example, during a PowerPoint™ slide presentation, a symbolic information capture device may record the slides by capturing slide transitions (e.g., by capturing keyboard commands) and then extracting the  
30 presentation images based on these transitions. Whiteboard capture devices may include devices such as a camera appropriately positioned to capture contents of the whiteboard, a screen, a chart, etc.

**[0033]** The information captured by capture devices 118 during a presentation may be stored in a repository or database 115 as recorded information 120. Recorded information 120 may be stored in various formats. For example, a directory may be created in repository 115 for storing recorded information 120, and the various types of information (e.g., audio information, video information, images, etc.) included in recorded information 120 may be stored in the directory. In another embodiment, recorded information 120 may be stored as a file. Various other techniques known to those skilled in the art may also be used for storing the recorded information. The image file may be referred to as being stored in a "document". A document is a file, directory, etc. that includes the image file.

**[0034]** One type of recorded information 120 may include captured image information. Captured image information includes one or more images captured by capture devices 118. For example, captured image information may be an image of a slide outputted from a symbolic presentation document. In one embodiment, a digital camera may have captured the image information. Temporal information, such as a time-stamp, may also be recorded with the captured image information. Captured image information may be stored in a "recorded images document".

**[0035]** Another type of recorded information 120 may be referred to as "recorded presentation information". The recorded presentation information includes information, such as audio and/or video information, that includes images of slides outputted from a symbolic presentation document. For example, recorded presentation information may be audio and/or video information of a presentation being given with slides outputted from a symbolic source document. Recorded presentation information may be stored in a "recorded presentation information document".

**[0036]** In addition to storing information recorded during a presentation, a symbolic presentation document may be stored in repository 115. It will be understood that the symbolic presentation document may also be stored in other repositories, such as in computer 102. The image file in repository 115 then may include recorded information 120 and symbolic presentation documents.

**[0037]** Server 112 is configured to store association information for stored information. The stored information may be referred to as an image file. The image file, for example, may include information from a recorded images document or symbolic presentation document. The association information that is stored associates an image file with a recorded

presentation information document. When a presentation is recorded, association information to a recorded presentation information document is stored for a document that includes images of slides outputted during the presentation. Also, temporal information that indicates when slides were outputted or captured may also be stored for the image file.

5    **[0038]**   The recorded presentation information document includes images of slides displayed on output device 116 during the presentation. As will be described in more detail below, an input image may be used to access portions of the recorded presentation information document. For example, an image of a slide being displayed on output device 116 may be used to determine an image file that includes information that is consider to  
10   match the image. In one embodiment, the image file may be a recorded images document or a symbolic presentation document. Association information for the image file is then used to determine a recorded presentation information document in a plurality of recorded presentation information documents that includes information that is considered to match the image. A portion of the recorded presentation information document that includes  
15   information that that is considered to match information in the image file is then determined.

**[0039]**   Fig. 2 depicts a simplified flow chart 200 for determining and storing association information for an image file according to one embodiment of the present invention. The method may be performed by software modules executed by a data processing system, hardware modules or combinations thereof. Flow chart 200 depicted in Fig. 2 is merely  
20   illustrative of an embodiment incorporated in the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize variations, modifications, and alternatives.

**[0040]**   In step 202, a slide from a symbolic presentation document is outputted. As shown in Fig. 1, a slide may be displayed on an output device 116.

25   **[0041]**   In step 204, information that includes an image of the outputted slide is captured. As described above, the image may be captured by capture devices 118. For example, the image may be captured by a digital camera. Also, the image of the slide may be captured by a presentation recorder or symbolic presentation device. It will also be understood that other methods may be used to capture an image of the outputted slide.

30   **[0042]**   The contents of the image may include information from the outputted slide. For example, the contents may include text and other semantic information. Also, color, layout, edges, and other features of an image may be derived from the input image to facilitate



comparison. For example, an edge histogram may be derived from an image of the outputted slide.

[0043] In step 206, the captured information of the outputted slide is then stored in one or more recorded images documents. In one embodiment, a series of slide images from a presentation may be stored in a recorded images document. For example, a series of slides that were outputted from a symbolic presentation document for a presentation may be stored in a file or directory in repository 115. In one embodiment, images of slides captured may be stored in a captured image information document, a recorded images document, and a recorded presentation information document.

[0044] In step 208, association information that associates an image file with the recorded presentation information document is stored. For example, information is stored so that the recorded images document or symbolic presentation document may be associated with the recorded presentation information document. The association information may associate an image file document to a recorded presentation document. Also, the association information may associate a portion of the image file document to a portion of the recorded presentation document. The association information may be stored in an XML document, which may be embedded in the recorded images document or stored separately.

[0045] In another embodiment, association information is stored that associates the symbolic presentation document that outputted the slide in step 202 with the recorded presentation information document. For example, information is stored so that the symbolic presentation document may be associated with the recorded presentation information document. The association information may be stored in an XML document, which may be embedded in the symbolic presentation document or stored separately.

[0046] In step 210, temporal information is stored for the image file. For example, temporal information may be stored for the recorded images document or symbolic presentation document. The temporal information may indicate a time that the image was captured. Also, the temporal information may indicate a time that the slide was output from a symbolic presentation document. For example, the temporal information may be a time-stamp. The temporal information may also be a range of times. For example, the range of times may be a time period that a slide is displayed during a presentation.

[0047] If times are synchronized where the desired point in the recorded presentation information document corresponds to a time that the outputted slide was captured, the

temporal information may be used to determine a portion in the recorded presentation information document that includes the image. The time of the capture device is synchronized with the time of a presentation capture device. For example, if an outputted slide is captured at 1:00 p.m., a point in the recorded presentation information document where the outputted slide may have been displayed is at 1:00 p.m. If the time was synchronized between the recorded presentation information document and capturing of the outputted slide, the point at 1:00 pm in the recorded presentation information document corresponds to when the slide was outputted and may include information that that is considered to match the image of the outputted slide.

[0048] The temporal information may also be stored as index points or access points. For example, a table may be created that associates a captured image with an index point or access point in the recorded presentation information document. For example, a pointer to a portion of the recorded presentation information document is stored for each captured slide. An XML table may be used to store the index or access points for the captured image.

[0049] Accordingly, using the association information, the image file, such as a symbolic presentation document or a recorded images document, may be associated with a recorded presentation information document. The temporal information may then be used to determine a portion of the recorded presentation information document. As will be described in more detail below, the association and temporal information is used to associate captured image information with a portion of the recorded presentation information.

[0050] Fig. 3 depicts a simplified flow chart 300 for a method of determining a portion of a recorded presentation information document according to one embodiment of the present invention. The method may be performed by software modules executed by data processing modules, by hardware modules or combinations thereof. Flowchart 300 depicted in Fig. 3 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention as recited in the claims, one of ordinary skill in the art would recognize variations, modifications and alternatives.

[0051] In step 302, an input image is determined. In one embodiment, the input image is an image from a recorded images document in recorded information 120. For example, the input image is an image taken by a digital camera.

[0052] In one embodiment, a user may input an identifier for an input image. For example, the identifier may include a storage location that stores the input image. Also, the input

image may be received from a device that is storing the image. For example, a capture device 118 that is storing the image may download the input image. Additionally, a hard copy of an image may be scanned by a scanner to generate the input image. A person of skill in the art will appreciate other methods for determining an input image.

5    **[0053]**   In step 304, an image file that includes information that that is considered to match the input image is determined. The input image may be compared to images extracted from the image file to determine an image file that includes information that that is considered to match the input image. For example, a symbolic presentation document or a recorded images document that includes information that that is considered to match the input image is  
10   determined.

**[0054]**   The information that that is considered to match the input image may include an image of a slide that was outputted on output device 116. A capture device 118 may have captured the image of the slide that was outputted on output device 116. The image may then include information that that is considered to match the input image.

15   **[0055]**   The information that matches the input image may also be an image of a slide in a symbolic presentation document. The input image once again may include an image of a slide outputted on output device 116. The slide was outputted using a slide in the symbolic presentation document. Thus, an image of the slide in symbolic presentation document includes information that matches the input image of a slide outputted on output device 116.

20   **[0056]**   Different techniques may be used to compare an input image to information in an image file. In one embodiment the information that matches the input image is determined using content matching techniques. Examples of content matching techniques that may be used are described in U.S. Application No. 10/412,757, Attorney Docket No. 15358-009500US, filed April 11, 2003, which is hereby incorporated by reference for all purposes.  
25   In one embodiment, an input image is compared to keyframes extracted from the image file document to determine matching information.

**[0057]**   In one embodiment, text may be extracted from the input image and compared with text in images extracted from the image file document. For example, text from images of the slide in the symbolic presentation document or the recorded images document may be  
30   extracted and compared to text extracted from the input image. An image in the image file and the input image may be considered to be matching when the extracted text substantially matches.

[0058] In step 306, using association information, a recorded presentation information document that is associated with the image file determined in step 304 is determined. In one embodiment, association information is determined for an image file. For example, an XML document or a table that lists recorded presentation information documents and which image file is associated with them may be used to determine associated recorded presentation information documents for the image file. In one embodiment, an image file includes images of slides that are outputted and recorded in a recorded presentation, that image file is associated with the recorded presentation information document.

[0059] The association information may be determined manually or automatically. For example, a user may manually associate the recorded presentation information document with the image file. Also, when a slide is outputted using a symbolic presentation document, information may be stored that associates a recorded presentation information document that includes information that is considered to be matching an image of the outputted slide in an image file. Further, content matching techniques may be used to determine contents of an image file that includes information that matches information in the recorded presentation information document. If information matches, then the recorded images document or symbolic presentation document that includes the matching information may be associated with the recorded presentation information document.

[0060] In step 308, a portion of the recorded presentation information document is determined. In one embodiment, the portion includes information that matches the input image. For example, the portion may include an image of the input image. Thus, information in a slide that is depicted in the input image may match information in an image of a slide in the recorded presentation information document.

[0061] In one embodiment, temporal information for the input image or image file may be used to determine the portion of the recorded presentation. For example, a time-stamp associated with the input image is used to determine the portion of the recorded presentation. If the recorded presentation and input image were synchronized, a time-stamp for the input image may be used to determine the portion in the recorded presentation information document. For example, if an input image was captured at 1:00 p.m., information in the recorded presentation information document at a time corresponding to 1:00 p.m. may match information in the input image.

**[0062]** A time range based on the time stamp may be used to determine the portion of the recorded presentation information document. For example, a certain amount of time before and after the time stamp may be used to determine the portion of the recorded presentation information document. The time range may have been stored with the input image and is used to determine the portion of the recorded presentation information document.

Alternatively, the time range may be determined based on the time-stamp. For example, a certain amount of time may be added or subtracted from the time-stamp to determine the time range. Also, the time range may correspond to a length of time that a slide was displayed. Accordingly, the time range may include time that is before or after the time that a slide was displayed or the time stamp used to determine the portion of the recorded presentation information document.

**[0063]** In another embodiment, information from the image file, such as an image of a slide in the symbolic presentation document or recorded images document, may be used to determine a portion of the recorded presentation. For example, the image of the slide in the recorded images document may include time information that indicates when the image of the slide was captured. Also, the image of the slide in the symbolic presentation document may include time information that indicates when the image of the slide was outputted. If the stored time information is synchronized with the recorded presentation information, then the time information may be used to determine the portion in the recorded presentation information document. As mentioned above, the time information may be a range of time.

**[0064]** In another embodiment, index or access points for the image file may be used to determine the portion in the recorded presentation information document. The index or access points may have been stored in an XML document and point to portions in the recorded presentation information document.

**[0065]** In another embodiment, content matching techniques may be used to determine the portion of the recorded presentation information document. For example, a time period may be determined where the input image was displayed in the recorded presentation information document. Text may be extracted from the input image and the recorded presentation information document and compared to determine a time period in the recorded presentation where the extracted text matches. For example, a slide that includes text information that matches the extracted text from the input image may be displayed for a time period. It will

be understood that the portion of the recorded presentation information document determined may also include time before or after the time period determined.

[0066] In another embodiment, the physical location of a device, such as capture device 118, may be used to determine the portion of the recorded presentation information document. For example, the device's outdoor location may be used to interrogate a mapping service to determine a location. In one example, an image of a menu captured by a user may be associated with restaurants in the vicinity of the user's current location. The device's indoor location may also be similarly used to determine the recorded presentation information document. For example, given that a device captured an image in The Explorer Room of the Fairmont Hotel, the image of a presentation slide may be matched only to presentations that occurred or were scheduled to occur in that room. An expected location for capture device 118 or for a user, such as provided from the appointment scheduler on the device, may be used in a similar way to constrain the matching process. The user's seat assignment may be used to determine a seating chart for the room. This may provide an expected distance and angle to the screen. These values may then be provided to the matching process, which may use them to determine the portion of the recorded presentation information document. For example, these values may be used to calculate the expected angular distortion of the text in the image and a correction process could be applied to the image before it is OCR'd. The angle of the device in 3-space, height of the ground and measured distance to the object may be used in a similar way to correct the image before it is recognized further.

[0067] In step 310, an action is performed with the portion of the recorded presentation information document. For example, the portion of the recorded presentation may be displayed in an interface. When the portion is displayed, a user may be given the option to play the portion of the recorded presentation information document. Also, the portion of the recorded presentation information document may be displayed and played without input from a user. Further, the portion of the recorded presentation information document may be sent to a device where the device may display the portion of the recorded presentation information document or play the portion recorded presentation information document.

[0068] Fig. 4 depicts a simplified flow chart 400 of an alternate embodiment for determining a portion in a recorded presentation information document using an input image according to embodiments of the present invention. The method may be performed by software modules executed by a data processing system, by hardware modules or

combinations thereof. Flow chart 400 depicted in Fig. 4 is merely illustrative of an embodiment incorporated in the present invention and does not limit the scope of the invention as recited in the claims. One of ordinary skill in the art will recognize variations, modifications, and alternatives.

5 [0069] In step 402, an input image is determined. In one embodiment, the input image includes an image of output device 116 and the outputted slide. For example, the input image may be captured by a digital camera that took a picture of output device 116 during a presentation.

10 [0070] In step 404, the contents of what the output device is outputting are determined. In one embodiment, the input image may include objects other than the outputted slide. For example, the image may include any surrounding objects located around output device 116, such as a picture of the presenter and any other objects around output device 116.

15 [0071] In determining the contents of what the output device is outputting, an image of the outputted slide may be determined. In one embodiment, the image of the outputted slide is determined by analyzing the image to determine an image of a slide in output device 116. Also, text of the slide may be extracted from the image of the slide. In one embodiment, the text of the slide that is found in the input image is determined. In another embodiment, other features of the display may be used to determine the image. For example, the color, layout, edges, and other features of the image or derived from the image may be used to determine  
20 the image. For example, an edge histogram may be derived from the image to determine an image of the outputted slide in the output device.

[0072] In step 406, a plurality of recorded presentation information documents are searched to identify a recorded presentation information document that includes the contents determined in step 404. In one embodiment, content matching techniques are used to  
25 compare the contents determined in step 404 to the contents of the plurality of recorded presentation information documents. Recorded presentation information documents may be associated with the input image when contents of a recorded presentation include information that matches the contents determined in step 404. For example, if a slide is output during a presentation, the slide may be included in the contents of a recorded presentation information  
30 document. Thus, if an input image of a slide is captured and matches contents of a recorded presentation information document, then that recorded presentation information document may be a presentation in which the input image was displayed. Although it is described that

one recorded presentation information document is determined, it will be recognized that multiple recorded presentation information documents may be determined. For example, the same slide may be captured in multiple presentations that are given.

**[0073]** In one embodiment, the text extracted from the contents of what an output device is displaying may be compared with text extracted from information in a recorded presentation information document. If the text substantially matches the text in an image of the recorded presentation information document, that recorded presentation information document is determined to include information that includes the contents determined in step 404.

**[0074]** In step 408, a portion in the recorded presentation information document is determined that includes the contents determined in step 404. In one embodiment, time information may be used to determine the portion in recorded presentation information document. For example, a time-stamp associated with the input image may be used to determine a portion in the recorded presentation information document. If the presentation and input image were captured at the same time, then the time stamp may indicate a portion in the recorded presentation information document where the input image was outputted. In this case, the recorded presentation information document may include the contents of output device 116 because the input image is being displayed. As described above, a portion of the recorded presentation information document may be a time range that may include time before or after the time stamp.

**[0075]** In another embodiment, as mentioned above, content matching techniques may be used to determine the portion of the recorded presentation information document. For example, a time period may be determined where the input image was displayed in the recorded presentation information document.

**[0076]** In another embodiment, as mentioned above, the physical location of a device, such as capture device 118, may be used to determine the portion of the recorded presentation information document.

**[0077]** In one embodiment, an action may be performed with the portion of the recorded presentation information document as described in step 310 of Fig. 3.

**[0078]** Fig. 5 depicts an interface 500 that may be used to specify an input image according to one embodiment of the present invention. As shown, an input image may be specified in an entry 502. For example, an identifier to a location where the image may be stored is input



in entry 502. The identifier may be a storage location, file identifier, URL or any other identifier. Also, a user may use a browse button 504 to browse for the input image among stored images.

[0079] In another embodiment, an input image is specified by entering keywords in entry 506. The keywords may specify content that is included in a slide image that is captured. For example, if a slide included the words "Multimedia Presentation", the keywords "Multimedia Presentation" may be entered to retrieve input images that include the keywords.

[0080] In one embodiment, multiple input images may be specified. The multiple images may be used to retrieve a presentation or a point in a presentation. In this case, multiple identifiers may be inputted into entry 502.

[0081] Fig. 6 depicts an interface 600 that displays a portions of the recorded presentation information document in response to receiving an input image according to one embodiment of the present invention. As shown, interface 600 includes a first window 602 that displays an input image. The input image acts as a query to find matching images. The query may have been specified by an identifier for an input image.

[0082] The input image is then retrieved and, as shown, is displayed in window 602 as input image 606. Input image 606 includes an image of a slide being displayed on an output device 116. Also, the image includes a picture of a presenter and also other objects surrounding the output device 116.

[0083] From input image 606, text has been extracted as shown in a section 608. The extracted text includes text that has been extracted from a slide that has been outputted on output device 116. The extracted text may be compared to information in an image file.

[0084] A second window 604 displays query results for the query. The query results include images of one or more slides determined from the image file and associated portions of a recorded presentation information document.

[0085] An image of a first slide is depicted in a window 610 and a first portion of the recorded presentation information document is depicted in window 612 and an image of a second slide is depicted in a window 614 and a second portion of the recorded presentation information document is depicted in a window 616. The image in window 610 includes information found in the recorded presentation information document depicted in window 612. For example, the image of a first slide in window 610 may be displayed in an output

device 116 found in the recorded presentation information document depicted in window 612. The same may true for the image of the second slide in window 614 and the second portion of the recorded presentation information document depicted in window 612. As shown, multiple slides may include information that matches the text extracted from the input image 606. In this case, slides may be different in that information other than text shown in the slides is different but the text is substantially the same. It will be understood that slides 610 and 614 may not include the same text as the extracted text 608 but may include substantially similar text.

**[0086]** A link 618 may also be included in window 604. Link 618 may point to other relevant image files. For example, link 618 may be associated with an agenda, a paper, etc. that is relevant to a slide depicted in window 610 or 614, or to the portion of the recorded presentation information document depicted in window 612 or 616.

**[0087]** For each portion of the recorded presentation information document, a match score 620 is displayed along with a slide duration 622. Match score 620 indicates the relevancy of a slide determined in the image file. For example, an algorithm may be used to calculate how relevant information in an image of a slide is to the extracted text. For example, the text extracted may be considered to match exactly with the text depicted in a slide. This may return a high match score. However, a slide may include more text or less text than the extracted text. An algorithm may determine how relevant the slide is. A user may then use match score 620 to determine which portion of the recorded presentation information document is most relevant.

**[0088]** A slide duration 622 shows a time length of the portion of the recorded presentation information document. For example, slide duration 622 may be how long the slide was displayed (plus or minus a certain amount of time). A user may have the choice to play the recorded presentation information document for the slide duration 622 or may choose to play additional parts of the recorded presentation information document before the beginning or ending point of the portion.

**[0089]** Fig. 7 depicts a simplified flow chart 700 of a method for using an input image to perform a search according to one embodiment of the present invention. The method may be performed by software modules executed by a data processing system, by hardware modules or combinations thereof. Flow chart 700 depicted in Fig. 7 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope of the invention

as recited in the claims. One of skill in the art would recognize variations, modifications, and alternatives.

[0090] In step 702, an input image is determined. The input image may be determined using techniques described above in step 302 of Fig. 3.

5 [0091] In step 704, features from the input image are extracted. For example, content, such as text, of a slide displayed in an output device 116 may be extracted.

[0092] In step 706, a search query from the extracted features is determined. In one embodiment, the text is analyzed to determine a search query. For example, words that appear more frequently than others may be used in a search query. Also, more commonly  
10 used words, such as "a", "of", etc., may be deleted from the extracted features and the remaining words used as a search query.

[0093] In step 708, a search is performed with the search query. For example, a search engine, such as Google, may be automatically invoked with the search query determined in step 706. Additionally, other data may be used in the search. For example, metadata  
15 specifying that only PowerPoint \*.ppt files should be returned may be included in the search query sent to search engine. In this case, \*.ppt files may be searched for on the Internet or Worldwide Web using an input image. Thus, any recorded presentation information documents that include text in a search query may be retrieved. Also, metadata may specify that only slide images (e.g., .smil files) should be returned.

20 [0094] In step 710, the results of the search are determined. The results may include \*.ppt files that include information that matches the search query. For example, images of slides that include text that is considered to match the search query are determined.

[0095] In one embodiment, the results may then be displayed for a user to choose a recorded presentation information document that is desired. Also, portions of a recorded  
25 presentation information document that includes information that matches the \*.ppt files may be determined. The portions may then be played, etc.

[0096] Fig. 8 depicts a simplified flowchart 800 of a method for determining television programs using an input image according to one embodiment of the present invention. The method may be performed by software modules executed by data processing system, by  
30 hardware modules or combinations thereof. Flowchart 800 depicted in Fig. 8 is merely illustrative of an embodiment incorporating the present invention and does not limit the scope

of the invention as recited in the claims. One of ordinary skill in the art would recognize variations, modifications, and alternatives.

[0097] In step 802, an input image including an image of a television screen is determined. In one embodiment, the image may be taken by a digital camera while a television program is being displayed. Also, a television recorder may be programmed to record an image of a television program at a certain point. The images may be determined using techniques described in step 306 of Fig. 3.

[0098] In step 804, the input image is compared to a plurality of television programs to determine a television program that includes information that matches information in the input image. In one embodiment, content based matching techniques described above may be used to determine the information that matches the input image. For example, the image of the television screen includes an image of a television program being displayed. The plurality of television programs are searched to determine a television program that includes the image that was captured. For example, the captured image may be matched to keyframes of the plurality of television programs.

[0099] Also, if the picture is taken when closed captioning is on, the text of the closed captioning may be extracted from the input image. In one embodiment, optical character recognition (OCR) or other techniques may be used to extract the text. The text is then used in comparing text extracted from the plurality of television programs to determine television programs that include information that matches the captured image.

[0100] In step 806, the television program determined in step 804 is retrieved. It will be recognized that multiple TV programs may be determined.

[0101] In step 808, a portion of the retrieved television program that includes the input image is determined. In one embodiment, the portion is determined using techniques described in step 308 of Fig. 3. For example, temporal information or content matching techniques are used to determine the portion. In one embodiment, the portion includes the image that is displayed on the television screen of the input image.

[0102] In step 810, an action is performed with the portion of the retrieved television program. For example, the television program may be displayed at the beginning point of the portion and a user may have the option of playing the portion. Also, the portion may be displayed and played. Also, the television program or portion of the television program may

be sent to a device or stored in a repository for later use by a user. The user may then view the portion of the television program at a later date.

[0103] Fig. 9 depicts an embodiment showing using a digital camera for retrieving a collection of recorded presentation information documents and determining a portion in the recorded presentation information documents according to embodiments of the present invention. As shown, a digital camera 1002 captures an image during a presentation. The input image of the presentation is shown in window 1004. The input image may include an image of a slide outputted by a symbolic presentation document. In one embodiment, the digital camera image may be different than a screen capture of an image, scan of the image, or image from a source document. A screen capture, scan and image from a source document include only a region of interest (ROI), such as the image of the slide. A digital camera image may include the ROI in addition to surrounding objects, such as the display device, presenter, etc. Also, the ROI from the digital camera image may be occluded and may include motion blur, which may not be included in a screen capture, scan, or image from a source document.

[0104] The image is used to retrieve recorded presentation information documents that include information that matches the input image. For example, recorded presentation information 1006-1, -2, -3, and -4 may be searched using techniques described above. In one embodiment, it is determined that recorded presentation information document 1006-2 includes information that matches the input image.

[0105] A portion of recorded presentation information document 1006-2 is then determined. The portion may be determined using techniques described above. The portion in recorded presentation information document 1006-2 may then be displayed in an interface 1008 and playback of the portion of recorded presentation information document 1006-2 is enabled.

[0106] Accordingly, a digital camera may be used to determine a portion of a recorded presentation information document in a plurality of recorded presentation information documents. Thus, when a user is interested in a certain slide that has been displayed in a presentation, a digital camera may be used to record the moment in which the slide is outputted. A time stamp may be attached to the image and used to determine a portion in the recorded presentation information document that includes an image of the slide that was displayed.

[0107] Fig. 10 depicts a simplified flowchart 900 of a method for performing a disambiguation process in determining a portion of one or more recorded presentation information documents according to one embodiment of the present invention. A device 902 is configured to capture an image. The image and device information is sent from device 902 and is received in step 904. The device information includes any information associated with device 902. For example, the device information may include an address for device 902, such as an identifier for device 902. The identifier may be a phone number if device 902 is cellular phone, a network address, a URL, an 802-11 identifier, etc.

[0108] Also, device information may include a command and parameters for the command. The command may indicate how a disambiguation process should be executed. For example, the command may indicate that a match and store process should be performed. In this case, if one or more recorded presentation information documents are determined to include information that matches the image, the one or more recorded presentation information documents should be stored without further feedback from device 902. Also, the command may indicate that an interactive query should be used, in which case information should be sent to device 902 in order for disambiguation of the results to be performed. If a "results only" command is used, parameters may be specified that indicate the type of result (e.g., PowerPoint file, translated PowerPoint file, business card, related document, streamed audio from a conference room, mpeg file for a movie playing in a theater, translated menu, etc.) and a number of results that should be returned. Also, if a command is an interactive query and results command, device 902 should be used to disambiguate the one or more recorded presentation information documents returned and also device 902 should receive the results of the disambiguation process.

[0109] In step 906, steps 302, 304, 306, 308 of Fig. 3 or steps 402, 404, 406, and 408 of Fig. 4 are performed to determine a portion of one or more recorded presentation information documents.

[0110] When portions in one or more recorded presentation information documents are determined, in step 908, it is determined if the results should be disambiguated. A command indicating whether the disambiguation should be performed may have been received with the device information received in step 904. Also, it may be determined from the results if disambiguation is needed. For example, it may be unclear which recorded presentation information documents include information that matches the input image and thus

disambiguation may be needed to determine the one or more recorded presentation information documents.

[0111] If disambiguation is not necessary, one or more recorded presentation information documents have been determined.

5 [0112] If disambiguation is desired, information is sent to device 902 for disambiguation. For example, device 902 may be sent images of portions of several recorded presentation information documents. A user may then be prompted to determine which recorded presentation information documents are desired. Additionally, questions may be asked that are used to determine recorded presentation information documents that should be retrieved.  
10 For example, a user may be prompted with the question, "Is this Mary Smith's presentation or Bob Jones' presentation?" Depending on the answer, a certain recorded presentation information document may be more applicable than another. Alternatively, a user may be asked to take another picture of the original object perhaps with some alteration in lighting, aspect ratio, or zoom. One reason a user may be prompted to take another picture is because  
15 the feature extraction (e.g., the text extraction produced poor results) or too many recorded presentations have been returned by the content base linking.

[0113] When the results of the disambiguation process are received from device 902, in step 910, disambiguation processing is performed. For example, if a user had selected certain recorded presentations, those recorded presentations are determined to be the desired ones.  
20 Also, additional searches may be performed with the information received from the results of the disambiguation process.

[0114] After the disambiguation process is performed, in step 908, it is determined if additional disambiguation processing should be performed. If not, one or more recorded presentation information documents are determined and an action may be performed.  
25 Actions may include retrieving information associated with a presentation such as a presenter's business card, the presenter's home page, documents from the presenter's publication list, etc. Other examples include when a picture of a sign on the front of a restaurant is taken, a menu from the restaurant's website may be returned; when a picture in front of a theater is taken, movie times for movies currently showing in the theater may be  
30 returned; when a picture of a document is taken, a translation of the document is returned; and when a picture of a bar code is taken, information for the item associated with the bar code is returned.

[0115] If disambiguation is still needed, the process described above continues.

[0116] The disambiguation process allows a user to communicate with a system to determine one or more portions of the one or more recorded presentation information documents. Also, the device information may be used to open a communication channel  
5 between the system and device 902. For example, a streaming data link between recording devices in a room and device 902 may be established when an input image is received. The picture of a presentation being given would indicate that a user is present in a location where a presentation being recorded is taking place. The system may send the recorded presentation information to device 902. The communication may continue until the presentation was  
10 complete or the user actively disconnects it. Accordingly, a user can receive an audio / video recording of a presentation while it is being given.

[0117] Fig. 11 is a simplified block diagram of a data processing system 1100 that may incorporate an embodiment of the present invention. As shown in Fig. 11, data processing system 1100 includes at least one processor 1102, which communicates with a number of  
15 peripheral devices via a bus subsystem 1104. These peripheral devices may include a storage subsystem 1106, comprising a memory subsystem 1108 and a file storage subsystem 1110, user interface input devices 1112, user interface output devices 1114, and a network interface subsystem 1116. The input and output devices allow user interaction with data processing system 1102.

[0118] Network interface subsystem 1116 provides an interface to other computer systems, networks, and storage resources 1104. The networks may include the Internet, a local area network (LAN), a wide area network (WAN), a wireless network, an intranet, a private network, a public network, a switched network, or any other suitable communication network. Network interface subsystem 1116 serves as an interface for receiving data from  
20 other sources and for transmitting data to other sources from data processing system 1100. Embodiments of network interface subsystem 1116 include an Ethernet card, a modem (telephone, satellite, cable, ISDN, etc.), (asynchronous) digital subscriber line (DSL) units, and the like.

[0119] User interface input devices 1112 may include a keyboard, pointing devices such as a  
30 mouse, trackball, touchpad, or graphics tablet, a scanner, a barcode scanner, a touchscreen incorporated into the display, audio input devices such as voice recognition systems, microphones, and other types of input devices. In general, use of the term "input device" is



intended to include all possible types of devices and ways to input information to data processing system 1100.

[0120] User interface output devices 1114 may include a display subsystem, a printer, a fax machine, or non-visual displays such as audio output devices. The display subsystem may be a cathode ray tube (CRT), a flat-panel device such as a liquid crystal display (LCD), or a projection device. In general, use of the term "output device" is intended to include all possible types of devices and ways to output information from data processing system 1100.

[0121] Storage subsystem 1106 may be configured to store the basic programming and data constructs that provide the functionality of the present invention. For example, according to an embodiment of the present invention, software modules implementing the functionality of the present invention may be stored in storage subsystem 1106. These software modules may be executed by processor(s) 1102. Storage subsystem 1106 may also provide a repository for storing data used in accordance with the present invention. For example, the images to be compared including the input image and the set of candidate images may be stored in storage subsystem 1106. Storage subsystem 1106 may comprise memory subsystem 1108 and file/disk storage subsystem 1110.

[0122] Memory subsystem 1108 may include a number of memories including a main random access memory (RAM) 1118 for storage of instructions and data during program execution and a read only memory (ROM) 1120 in which fixed instructions are stored. File storage subsystem 1110 provides persistent (non-volatile) storage for program and data files, and may include a hard disk drive, a floppy disk drive along with associated removable media, a Compact Disk Read Only Memory (CD-ROM) drive, an optical drive, removable media cartridges, and other like storage media.

[0123] Bus subsystem 1104 provides a mechanism for letting the various components and subsystems of data processing system 1102 communicate with each other as intended. Although bus subsystem 1104 is shown schematically as a single bus, alternative embodiments of the bus subsystem may utilize multiple busses.

[0124] Data processing system 1100 can be of varying types including a personal computer, a portable computer, a workstation, a network computer, a mainframe, a kiosk, or any other data processing system. Due to the ever-changing nature of computers and networks, the description of data processing system 1100 depicted in Fig. 11 is intended only as a specific example for purposes of illustrating the preferred embodiment of the computer system. Many

other configurations having more or fewer components than the system depicted in Fig. 11 are possible.

[0125] While the present invention has been described using a particular combination of hardware and software implemented in the form of control logic, it should be recognized that other combinations of hardware and software are also within the scope of the present invention. The present invention may be implemented only in hardware, or only in software, or using combinations thereof.

[0126] The above description is illustrative but not restrictive. Many variations of the invention will become apparent to those skilled in the art upon review of the disclosure. The scope of the invention should, therefore, be determined not with reference to the above description, but instead should be determined with reference to the pending claims along with their full scope or equivalents.